



Universidade Federal do Rio de Janeiro  
Centro de Ciências Jurídicas e Econômicas  
Faculdade de Administração e Ciências Contábeis  
Curso de Biblioteconomia e Gestão de Unidades de Informação



ROBSON SOARES CRUZ

## DESENVOLVIMENTO DE UMA POLÍTICA DE INDEXAÇÃO PARA O OBSERVATÓRIO DE GESTÃO DA INFORMAÇÃO

Rio de Janeiro  
2011  
ROBSON SOARES CRUZ

# DESENVOLVIMENTO DE UMA POLÍTICA DE INDEXAÇÃO PARA O OBSERVATÓRIO DE GESTÃO DA INFORMAÇÃO

Trabalho de conclusão de Curso  
apresentado ao Curso de Biblioteconomia e  
Gestão de Unidades de Informação  
(CBG/FACC), da Universidade Federal do  
Rio de Janeiro, como requisito parcial para  
obtenção do Grau de Bacharel em  
Biblioteconomia.

Orientador(a): Profa Dra Vânia Lisboa da Silveira Guedes  
Coorientador(a): Profa Dra Ana Maria Barcellos Malin

Rio de Janeiro  
2011

C955d Cruz, Robson Soares.  
Desenvolvimento de uma política de indexação para o  
Observatório de Gestão da Informação / Robson Soares Cruz. –  
Rio de Janeiro, 2011.

34 f. : il.

Orientadora: Vânia Lisboa da Silveira Guedes / Coorientadora:  
Professora Ana Maria Barcellos Malin.

Trabalho de Conclusão de Curso (Graduação em Biblioteconomia)  
– Curso de Biblioteconomia e Gestão de Unidades de Informação,  
Universidade Federal do Rio de Janeiro.

1. Política de indexação. 2. Análise de assunto. 3. Indexação  
temática. 4. Metadados.

CDD: 020

ROBSON SOARES CRUZ

DESENVOLVIMENTO DE UMA POLÍTICA DE INDEXAÇÃO PARA O  
OBSERVATÓRIO DE GESTÃO DA INFORMAÇÃO

Trabalho de conclusão de Curso  
apresentado ao Curso de Biblioteconomia e  
Gestão de Unidades de Informação  
(CBG/FACC), da Universidade Federal do  
Rio de Janeiro, como requisito parcial para  
obtenção do Grau de Bacharel em  
Biblioteconomia.

BANCA EXAMINADORA:

Aprovado em:

---

Prof.<sup>a</sup> Vânia Lisboa da Silveira Guedes – UFRJ  
Doutora em Linguística  
**Orientador(a)**

---

Prof.<sup>a</sup> Ana Maria Barcellos Malin – UFRJ  
Doutora em Ciência da Informação  
**Coorientador(a)**

---

Prof.<sup>a</sup> Maria José Veloso da Costa Santos – UFRJ  
Mestre em Ciência da Informação  
**Professora Convidada**

---

Prof.<sup>a</sup> Maria Irene da Fonseca e Sá – UFRJ  
Mestre em Engenharia de Sistemas e Computação  
**Professora Convidada**

Dedico este trabalho a minha família,  
pelo apoio constante em todos os momentos.

## **AGRADECIMENTOS**

Agradeço a meus pais por estarem sempre presentes e atenciosos. Por terem creditado sua fé em mim e nesta jornada que iniciei, me apoiando com positivismo e companhia.

A meu irmão por entender as dificuldades que vivi e apoiar todas as minhas escolhas.

A minha querida namorada Taís, pelo amor, carinho, força e companhia, me auxiliando nas decisões e dificuldades.

A meus amigos pelos momentos de diversão e alívio de todo o estresse.

A professora Vânia Guedes pelo constante apoio e auxílio neste trabalho.

A professora Ana Malin, por ter me aceito no projeto que dá origem a este trabalho, sendo uma grande inspiração para mim por seus conhecimentos, sabedoria, e pelo apoio.

*“Nada é tão real como um sonho e, se você for atrás dele, algo maravilhoso lhe acontecerá:  
pode ser que você envelheça, mas jamais será velho.”*

Tom Clancy

## RESUMO

CRUZ, Robson Soares Cruz. **Desenvolvimento de uma política de indexação para o Observatório de Gestão da Informação**. 2011. 34f. Trabalho de Conclusão de Curso (graduação). Curso de Biblioteconomia e Gestão de Unidades de Informação. Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2011.

O presente trabalho propõe o desenvolvimento de uma política de indexação para o Observatório de Gestão da Informação. A pesquisa insere-se no âmbito da linha de estudos voltados para a indexação das informações postadas nesse *site*, dando sequência a projetos recentes voltados para a área de Gestão da Informação e do Conhecimento no Curso de Biblioteconomia e Gestão de Unidades da Informação da Universidade Federal do Rio de Janeiro. Observa-se inicialmente a utilidade da indexação de informações recuperadas através de mecanismos automáticos de busca na *web*, que utilizam metadados para a representação temática e organização da informação. Percebe-se que nesse ambiente, interligado pela rede mundial de computadores, há um crescimento exponencial e desordenado de páginas e, sobretudo, de seus conteúdos e representações. A problemática da recuperação de informação neste ambiente está na dificuldade especialmente de se obter consistência na atribuição de termos de indexação e organização do conhecimento para que sistemas de busca recuperem informações precisas e relevantes. Através do desenvolvimento e implementação de uma política de indexação, que norteie a análise de assunto e a representação das informações que circulam em ambiente *web*, é demonstrado como os termos extraídos da linguagem natural de conteúdos nesse ambiente possibilitam a otimização da indexação temática e, conseqüentemente, a recuperação de informações relevantes em páginas dessa natureza.

Palavras Chave: Política de Indexação. Análise de assunto. Indexação temática. Metadados.



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	10
<b>2</b>	<b>JUSTIFICATIVA</b>	12
<b>3</b>	<b>OBJETIVOS</b>	13
3.1	Objetivo Geral	13
3.2	Objetivo Específico	13
<b>4</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	14
4.1	O Processo de Indexação	14
4.2	Análise de assunto	16
4.2.1	Leitura documentária	17
4.3	Indexação na <i>web</i>	18
4.4	A Política de indexação	20
4.5	Observatório de Gestão da Informação	22
<b>5</b>	<b>METODOLOGIA</b>	23
<b>6</b>	<b>RESULTADOS</b>	26
<b>7</b>	<b>CONSIDERAÇÕES</b>	32
	<b>REFERÊNCIAS</b>	33

## 1 INTRODUÇÃO

A explosão quantitativa da informação é um fenômeno que vem crescendo e atingindo patamares cada vez mais nebulosos. A produção da informação vem se intensificando em diversos meios e nem sempre essas informações são importantes para certos indivíduos com necessidades específicas.

De acordo com Le Coadic (2002, p.5), a explosão da quantidade de informação fundamenta-se num processo evolutivo, técnico e social do poder da linguagem e do raciocínio lógico. O autor explica que, com o advento da escrita, a comunicação passou de oral a escrita. Com o tempo a informação se multiplicou, através de cópias de manuscritos, fotocópias, impressões e passaram a ser armazenadas primeiramente em bibliotecas.

A multiplicação e crescente armazenamento de informações caracterizaram a “explosão de informação”. Com o advento da digitalização e depois com a *internet*, o processo de multiplicação da informação se intensificou. O crescimento da quantidade de informação que antes se deparava com barreiras físicas, agora tem um universo imaterial como fronteira.

A grande questão está na recuperação da informação. Diversos estudiosos hoje concentram esforços na tentativa de compreender como informação é organizada e recuperada em ambiente *web*. Para Rocha (2004, p.110) a dificuldade está na baixa organização da informação na *web*, que impede a construção de estratégias e mecanismos de busca eficientes.

Nesse ambiente, a indexação da informação como uma disciplina e técnica, pode ser utilizada de forma a permitir a categorização e classificação de conteúdos na *web*, representados por termos de indexação organizados em índices específicos, de forma que possibilitem a esses buscadores atingir um maior nível de precisão na recuperação da informação. De acordo com Lancaster (2004, p. 4) a precisão<sup>1</sup> na recuperação da informação está diretamente ligada à capacidade de evitar informações inúteis a uma necessidade de informação.

Esse estudo tem como objetivo principal propor o desenvolvimento de uma política de indexação que norteie a análise conceitual e a indexação temática das informações plotadas no

---

<sup>1</sup> A precisão é a capacidade do sistema de evitar documentos inúteis na recuperação de informação.

site do Observatório de Gestão da Informação (OGI), visando a contribuir para o estabelecimento de procedimento padronizado na indexação da informação de seus conteúdos informacionais e, conseqüentemente, à identificação de tendências relevantes em monitoramentos dessa natureza.

O trabalho foi dividido com tópicos e subtópicos para a melhor compreensão. O primeiro introduz de forma panorâmica o trabalho. O segundo apresenta uma breve justificativa para a escolha do tema. O terceiro item explicita o objetivo geral e os objetivos específicos pretendidos. O quarto apresenta a fundamentação teórica, considerando o processo de indexação, a análise de assunto, a leitura documentária, a indexação na *web*, a política de indexação e finalmente o Observatório de Gestão da Informação. O quinto item descreve a metodologia. O sexto apresenta os resultados, seguido das considerações finais. Por fim, são mencionadas as referências consultadas.

## 2 JUSTIFICATIVA

A importância do desenvolvimento de uma política de indexação, em qualquer sistema de informação, é sobretudo de permitir o estabelecimento de parâmetros que nortearão o processo de análise conceitual e atribuição de termos de indexação. Nesse sentido, a análise da informação e o processo de indexação devem ser embasados em uma política concisa, em consonância com os objetivos pré-determinados pelo sistema, com o intuito de atender à necessidade de informação do usuário que depende das terminologias adotadas pelo sistema para representar sua necessidade informacional.

Vale ressaltar que o processo, em si, de indexação pode ser desenvolvido de forma totalmente subjetiva se não se limitar a um objetivo que envolva o sistema, a terminologia e o usuário. Para Lucas (1999 apud JANNUZZI et al., 2005) o trabalho do indexador consiste em promover uma correlação entre a terminologia do documento analisado, a terminologia do sistema de informação e a terminologia do usuário.

Face à percepção de tais dificuldades, ao propor o desenvolvimento de uma política de indexação para o OGI objetiva-se estabelecer procedimentos e escolhas que visem a reduzir a subjetividade inerente ao processo de indexação, possibilitando a representação dos conteúdos dos textos presentes neste ambiente especialmente voltado para as necessidades dos usuários.

Em ambiente no qual o acesso à informação é de importância crucial para as discussões acadêmicas acerca da Gestão da Informação e do Conhecimento, os processos de indexação e organização da informação devem ser balizados por políticas de indexação que possibilitarão a constituição de estratégias de busca e recuperação da informação de acordo com os objetivos do sistema.

### 3 OBJETIVOS

Os objetivos deste trabalho estruturam-se em geral e específicos e são explicitados a seguir:

#### 3.1 Objetivo Geral

Desenvolver uma política de indexação, para o OGI, que vise à redução da subjetividade inerente ao processo de indexação temática, contribuindo para o refinamento do processo e organização do conhecimento sobre os temas discutido nas notícias e servindo de modelo para futuros indexadores neste ambiente.

#### 3.2 Objetivos Específicos

- Aumentar o nível de qualificação dos indicadores de tendências monitoradas pelo Observatório de Gestão da Informação, na área de Gestão da Informação e Gestão do Conhecimento.
- Viabilizar maior eficiência na busca e acesso à informação no sistema em análise.
- Comparar os resultados obtidos pela aplicação da Política de Indexação estabelecida no primeiro semestre de 2011 com os resultados obtidos no mesmo período do ano anterior, no qual não havia padronização e sistematicidade na análise conceitual e seleção dos termos de indexação.

## 4 FUNDAMENTAÇÃO TEÓRICA

Para a criação e o desenvolvimento de uma política de indexação no OGI, faz-se necessário o entendimento de alguns conceitos que norteiam os campos de indexação, com seus princípios de análise, síntese e representação da informação; do OGI e sua importância, assim como da política de indexação e suas aplicações, que no devido estudo visam ao ambiente *web*.

### 4.1 O Processo de Indexação

A indexação é definida como a forma de caracterizar os conceitos que estão presentes na escrita de certo tipo de documento de maneira a permitir a sua recuperação. O propósito principal da indexação é o de representar tematicamente esses documentos, através de termos de indexação, organizados em índices, que podem também ser denominados descritores, palavras-chave ou metadados. Segundo Méndez Rodríguez & Moreira González (1999 apud GUEDES, 2009, p. 3):

A indexação é uma técnica de classificação e caracterização do conteúdo, tanto do documento quanto das consultas formuladas pelos usuários, retendo as idéias mais representativas e vinculando-as a termos de indexação, extraídos da linguagem natural, empregada pelos autores, ou de um vocabulário controlado, selecionado a priori.

O processo de caracterização de documentos pela indexação não diz respeito à descrição física do material, como por exemplo, seu número de páginas, língua, data, formato, entre outros. Guedes (2009) menciona que o processo de indexação temática consiste na análise, síntese e representação da informação, por termos de indexação selecionados da linguagem natural do texto ou de uma linguagem documentária. Porém, de acordo com o *United Nations International Scientific Information System* (1976 apud PINTO, 1981, p. 84) os dados de descrição física podem ser adotados por indexadores que considerem essas informações úteis ao usuário no processo de recuperação.

Na indexação, o conteúdo do documento deve ser representado da forma mais precisa possível, através de descritores ou palavras-chave. Esse processo envolve subjetividade e exige o conhecimento do assunto e do contexto da comunidade assistida, assim como dos objetivos da organização. Após a etapa de indexação, os registros de informação são

organizados como índices, em livros, catálogos etc, como também em sistemas de metadados em ambiente *web*. Guedes (2009, p. 3) infere que o “índice, produto da indexação, consiste em uma relação de palavras ordenadas alfabeticamente ou sistematicamente, que indica a localização das informações” contidas no texto do documento em análise.

O processo de indexação acontece em duas etapas principais: a análise conceitual e a tradução. Essas duas etapas geralmente ocorrem de modo simultâneo. Segundo Lancaster (2004, p. 9), a análise conceitual, em primeiro lugar, implica em decidir sobre o que trata um documento, isto é, qual o seu assunto. Já a tradução seria a segunda etapa da indexação que implica em converter a análise conceitual processada num documento em um conjunto de termos de indexação.

Podemos salientar que não existe um padrão geral a seguir para se extrair termos de indexação de um documento. O mesmo documento pode ser indexado diferentemente em unidades de informação distintas. Este processo subjetivo é desempenhado por indexadores com níveis de conhecimento distintos, o que pode tornar a recuperação da informação complexa.

Existem ainda perguntas que o indexador deve fazer para nortear seu exercício: De que trata tal documento? Por que esse documento passou a fazer parte do acervo? Qual a importância desse documento para os usuários? Dependendo do objetivo, ponto de vista e foco nos usuários ou comunidade discursiva, a indexação abrangerá alguns aspectos em detrimento de outros e vice-versa. Como qualquer documento pode propiciar inúmeras visões e focos, a análise de assunto deve basear-se na clientela específica que é atendida. Hjørland (2001 apud LANCASTER, 2004, p. 10) acrescenta que:

Uma vez que qualquer documento pode, em princípio, proporcionar respostas a uma infinidade de questões, as análises de assuntos devem estabelecer prioridades baseadas nos grupos de usuários específicos atendidos (ou serviços específicos proporcionados na ecologia da informação). O assunto de um documento é assim relativo ao objetivo do serviço de informação específico.

Conforme mencionado, a segunda etapa da indexação é definida como Tradução. Existe aí uma distinção entre indexação derivada e indexação por atribuição. Na indexação derivada as próprias palavras que ocorrem no texto são selecionadas para representar seu conteúdo. Na

indexação por atribuição geralmente os termos são representados por um vocabulário controlado. Segundo Lancaster (2004, p. 19):

Um vocabulário controlado é essencialmente uma lista de termos autorizados. Em geral, o indexador somente pode atribuir a um documento termos que constem da lista adotada pela instituição para a qual trabalha. Comumente, no entanto, o vocabulário controlado é mais do que uma mera lista. Inclui, em geral, uma forma de estrutura semântica.

#### 4.2 Análise de assunto

Quando surge a necessidade de organização de acervos de maneira que permita a recuperação precisa de seus conteúdos, a representação documentária descritiva e temática permite que esse processo seja realizado de uma forma mais rápida e fácil. Na representação temática, para o indexador chegar à representação específica dos conceitos, que permitirá ao documento ser recuperado, é necessário efetuar a análise do assunto discutido no documento. A análise de assunto é um processo que permite identificar os conceitos abordados no documento e representá-los de forma específica e efetiva.

Para Fujita (2003, p. 61) a indexação como processo de análise documentária vem sendo realizada mais intensamente desde o aumento de publicações periódicas e da literatura técnico científica, criando a necessidade de mecanismos de controle bibliográficos em centros de documentação especializados. A autora explica que esse cenário, de crescente necessidade de representação de documentos, impulsionou o surgimento de mecanismos de controle bibliográfico, que não estavam exatamente vinculados a bibliotecas tradicionais.

O processo de análise do assunto, portanto, contribuirá significativamente para o estabelecimento de critérios para se identificar os conceitos mais representativos de cada documento na indexação. Cabe ao indexador a capacidade de analisar os conceitos abordados nos documentos para assim poder representá-los nos índices. A identificação dos conceitos é feita durante a leitura dos documentos.



#### 4.2.1 Leitura documentária

A análise de assunto possui a etapa da leitura como um momento muito importante, em que se identificam os conceitos de um documento, influenciando assim a sua indexação. Para Fujita (2004, p. 2) este processo é complexo porque depende do processamento humano de informações, da cognição de quem lê, do texto elaborado por um autor e do contexto de ambos. Assim essas diversas características influenciam no comportamento do indexador durante a leitura documentária.

Para Giasson (1993) e Cavalcanti (1989), citados por Fujita (2004, p. 2), esse processo de leitura pode ser entendido pela interação de três partes: texto, leitor e contexto. O texto possui os objetivos do autor e seu próprio contexto, assim como o leitor que também está em um contexto e possui seus próprios objetivos. Nesse sentido Fujita (2004) menciona que esse fenômeno:

(...) verifica-se, em situações de pesquisa formal e informal, obtenção de informação utilitária, formação profissional ou cultural ou de desempenho profissional, em que qualquer indivíduo, que se proponha a ler os textos existentes sobre o assunto pretendido, deverá realizar a leitura conforme objetivos de cada situação.”

A leitura documentária deverá ser processada de acordo com os objetivos da indexação, que terá como uma de suas finalidades a disseminação da informação para o usuário. Para Cavalcanti (1989 apud FUJITA, 2004, p. 5), a leitura possui esse aspecto de contexto e objetivos os quais o indexador deve estar atento, mas este também possui conhecimentos prévios, experiências acumuladas e valores, e utiliza essa carga de conhecimento para interagir com o texto. Essa interação entre texto e leitor possui essas características específicas que vão além do que o modelo Serial de Gough (1972 apud FUJITA, 2004, p. 5) apresentou, segundo o qual a leitura começa com uma fixação ocular sobre o texto, da esquerda para a direita de forma linear. Como salienta Paz (2006, p. 2), o modelo de leitura linear de Gough, passa a gerar muitas controvérsias junto a outros autores por também dizer que todas as letras devem ser consideradas individualmente no campo visual pelo leitor, antes de considerar qualquer série de letras. Para Goodman (1976, 1985 apud PAZ, 2006, p. 3) as letras não são a matéria da qual as sentenças são feitas. As letras devem estar associadas entre si para terem significado para quem as lê.

Esses diversos aspectos da leitura somam-se à característica comunicativa que este ato possui. O leitor não é mais um indivíduo passivo que recebe as informações e as digere. No ato da leitura há um processo de comunicação que se estrutura pelo emissor, mensagem, receptor e canal. Para Fujita (2004, p. 4) a informação, codificada e transmitida pelo emissor, passa através de um canal de comunicação e depois é decodificada pelo receptor, que devem possuir um código em comum, para que quando houver o “*feedback*” da informação não haja “ruído” por excesso de informação, nem “silêncio” por falta de informação. De acordo com a autora, a comunicação humana somente será durável se esta puder ser registrada em suportes documentários. Na perspectiva da comunicação, a leitura é um processo em que o autor é o emissor; o texto é a mensagem; o leitor é o receptor; a alfabetização é o canal comum; a compreensão é o “*feedback*”, e o que dificulta é o desconhecimento da língua, vocabulário ou assunto, por exemplo.

#### 4.3 Indexação na *web*

O processo de indexação se caracteriza pela seleção e atribuição de termos que representam o conteúdo temático de um documento. Porém esses documentos podem estar também na *internet*. Esses termos não são simples palavras como explica Rocha (2004, p. 111), são nomes dados a conceitos estabelecidos por uma comunidade e constituem o seu vocabulário controlado.

No entanto a grande rede de computadores torna esse processo de indexação e organização das informações mais complexo. Para Souza (2000 apud ROCHA, 2004, p.111) a *internet* difere de uma biblioteca tradicional pelo fato de não selecionar nenhum documento, pois abrange todas as áreas do conhecimento e seus usuários são todas as pessoas que a acessam.

A indexação de documentos em um acervo de biblioteca é baseada em uma política de indexação e indexadores voltados para as necessidades dos usuários que frequentam esse ambiente. Para um ambiente em que todos podem ser usuários da informação é preciso um processo de filtragem da informação. Para Rocha (2004, p. 112) a *internet* possui muitos documentos com conteúdos superficiais, redundantes e dúbios, assim como usuários que não têm uma conceituação muito clara do que querem obter em suas buscas, o que para o autor leva ao questionamento da eficiência da busca das informações indexadas por termos.

Outra questão complicada para a indexação na *internet* é o fato de que a rede possui uma infinidade de páginas, o que demandaria um esforço humano gigantesco para indexá-las. Para melhor representação dos conteúdos das páginas na *web* são utilizados metadados. Esse tipo de recurso pode ser entendido simplesmente por informações sobre informações ou dados sobre dados. Rocha (2004, p. 113) acrescenta que:

Quando se trata do mundo digital, chama-se de recurso o objeto descrito por metadados, pois este pode ser tanto um simples dado, quanto um documento, uma página da *web*, ou até mesmo uma pessoa, uma coleção, um sistema, um equipamento ou uma organização.

De acordo com o autor esses recursos significam qualquer objeto que pode ser acessado por meio de um endereço eletrônico, podendo ser páginas pessoais, *sites* ou sistemas. Os metadados representam os conteúdos de páginas na *internet* com a finalidade de facilitar aos recursos a sua localização. Para Feitosa (2006, p. 47) a *web* pode ter alguns mecanismos que auxiliam a recuperação da informação, como os metadados. As *tags* funcionam como metadados que representam os conteúdos das notícias.

Cleveland e Cleveland (2001 apud LANCASTER, 2004, p. 345) definem metadado como “dados sobre dados” que são estruturados para descrever recursos informacionais. Os autores explicam que bibliotecários e outros profissionais da informação podem criar metadados, os quais representam os assuntos dos documentos, assim como também a informação em meio *web*. Para Hock (2001 apud LANCASTER, 2004, p. 345), no campo da decodificação HTML de uma página na *web*, pode-se inserir metadados, o que ele mesmo chama de metaetiquetas, que funcionam como sinalizadores do conteúdo da página.

O princípio de indexação na *web* pode ser constituído por metadados na representação exhaustiva de conceitos abordados em páginas da mesma forma que descritores representam tematicamente um documento físico. Quanto mais representado por termos estiver um documento, ou seja, quanto mais exhaustivamente um sistema indexar este documento, maior será a sua revocação. Para Lancaster (2004, p. 4) a revocação é a capacidade do sistema para recuperar documentos úteis, e é inversamente proporcional, à precisão, que se refere a capacidade de evitar documentos inúteis.

#### 4.4 A Política de Indexação

A importância crucial da indexação para os sistemas de representação e recuperação da informação está pautada principalmente na elaboração de uma política de indexação que priorize os objetivos da organização na qual o sistema está inserido. Para Fujita e Rubi (2006, p. 49) esta política deve ser entendida como uma filosofia pertinente aos objetivos de recuperação da informação.

A política de indexação auxilia aos indexadores com parâmetros que diminuem a subjetividade e as incertezas, exercendo a função de um guia para tomada de decisão do processo como um todo. Para Carneiro (1985, p. 221 apud FUJITA; RUBI, 2006, p. 50):

[...] deve servir como um guia para tomada de decisões, deve levar em conta os seguintes fatores: características e objetivos da organização, determinantes do tipo de serviço a ser oferecido; identificação dos usuários, para atendimento de suas necessidades de informação e recursos humanos, materiais e financeiros, que delimitam o funcionamento de um sistema de recuperação de informações.

Para Kobashi (1994, p.17-19 apud FUJITA; RUBI, 2006, p. 51) os principais critérios da política de indexação devem ser condicionados às características dos sistemas documentários, como:

- necessidades do usuário;
- instituição onde se desenvolve;
- domínio tratado;
- recursos humanos, físicos e financeiros disponíveis;
- produtos e serviços;
- relação custo/desempenho.

Essas características apresentadas pela autora se referem a diversos modelos de aplicação da política de indexação. No que diz respeito ao OGI, considerou-se mais pertinente a aplicação dos critérios definidos não a partir de uma sequência de parâmetros primariamente definidos, mas sim como uma filosofia a ser aplicada ao modelo de busca da informação utilizado no OGI.

Para a aplicação dessa política de indexação, decidiu-se adaptar os elementos apresentados por Carneiro (1985 apud FUJITA; RUBI, 2006, p. 51-52), Robredo (1986, p. 80 apud FUJITA; SILVA, 2004, p.145) e Van Rijsbergen (1979 apud GUEDES, 2009) a fim de conciliar os objetivos do OGI com as necessidades de desenvolvimento de uma indexação em ambiente *web*. Assim os elementos escolhidos para compor a Política de Indexação do OGI, foram:

- a) Identificação da organização a qual está vinculado o sistema de indexação (contexto).
- b) Identificação da clientela a que se destina o sistema (destinatário).
- c) Cobertura de assuntos: centrais e periféricos abordados pelo sistema.
- d) Tipo de documentos.
- e) Línguas.
- f) Período de tempo: data de cobertura dos assuntos.
- g) Processo de indexação:  
Nível de exaustividade.  
Nível de especificidade da linguagem de indexação.  
Capacidade de revocação e precisão do sistema.  
Linguagem de Indexação: Linguagem natural ou linguagem documentária.  
Tipo de indexação: Indexação pós ou pré-coordenada.
- h) Avaliação do sistema.

É importante salientar que esse modelo apresentado por Carneiro (1985 apud FUJITA; RUBI, 2006, p. 51-52), Robredo (1986, p. 80 apud FUJITA; SILVA, 2004, p.145) e Van Rijsbergen (1979 apud GUEDES, 2009) refere-se à realidade dos sistemas de informação quanto a cobertura de assunto, seleção e aquisição dos documentos, sendo que no presente trabalho se apresenta adaptado de acordo com os objetivos do OGI de se monitorar tendências em Gestão da Informação (GI) e Gestão do Conhecimento (GC).

O estabelecimento destes critérios para a política de indexação permitirá que a relação entre o usuário do sistema, no caso dos conteúdos da *web*, e os documentos seja interativa, onde os dois lados saem ganhando. Pretende-se assim obter um modelo de indexação que possa substituir o anterior, o qual não é pautado em um padrão, tornando-se, portanto, uma referência para indexadores que futuramente se encontrem engajados no Projeto OGI.

#### 4.5 Observatório de Gestão da Informação

O modelo de política de indexação proposto será aplicado ao Observatório de Gestão da Informação<sup>2</sup>, que tem como finalidade servir como posto de observação na *web* para monitorar tendências relacionadas aos campos da Gestão da Informação e do Conhecimento (GIC), no Brasil.

O OGI foi criado e desenvolvido sob a coordenação e orientação da professora Doutora Ana Maria Barcellos Malin, contando com a participação dos alunos do Curso de Biblioteconomia e Gestão de Unidades de Informação da Universidade Federal do Rio de Janeiro. A questão inicial estava em como pesquisar sobre GI e GC, áreas em crescimento e ainda em formação, e assim desenvolver um projeto com o objetivo de monitoração de informações nessas áreas do conhecimento.

Através da utilização do mecanismo automático de busca “Alerta do Google” foram cadastrados os termos “Gestão da Informação” e “Gestão do Conhecimento”. Com esses registros os alunos selecionam as notícias de maior relevância de acordo com a finalidade do OGI, as categorizam e indexam, utilizando termos de indexação mais relevantes ao tema.

O processo de escolha das notícias se dá mediante uma filtragem realizada também pelo recurso “Alerta do Google”, no qual são feitos registros com as expressões já citadas, e os alertas enviam as notícias da *web* que seguem o modelo. A partir da seleção e categorização de cada notícia, elas são inseridas no portal *Wordpress* onde se localiza a página do OGI.

Nas etapas posteriores, ocorrem a análise e seleção dos termos mais significativos dessas notícias. Através da leitura e interpretação dos assuntos mais relevantes para o OGI, são selecionadas as palavras-chave que irão compilar uma “nuvem de *tags*” disponibilizada pelo sistema do *Wordpress*, permitindo que ocorram buscas através dos termos selecionados. Como esse processo é feito através da análise e tradução dos assuntos pelos indexadores e baseado em suas interpretações, conhecimentos e parcialidades, a Política de Indexação

---

<sup>2</sup> Disponível em <http://ogimonitorandoagora.wordpress.com/>.

proposta neste trabalho visa justamente a contribuir com o desenvolvimento de um parâmetro para que a atividade de indexação no OGI seja menos subjetiva e mais efetiva.

## **5 METODOLOGIA**

A metodologia adotada fundamenta-se no levantamento da literatura específica sobre o tema indexação temática e política de indexação, que forneceram embasamento teórico e metodológico para a estruturação de uma política de indexação voltada para os usuários do OGI.

Esta política foi definida de acordo com a finalidade do OGI que consiste no monitoramento de notícias e assuntos relativos à Gestão da Informação e Gestão do Conhecimento, com ênfase para os aspectos que permitem o aumento da qualificação dos indicadores de tendências já delineados. Assim foram propostos elementos para a otimização do processo de indexação que tem por objetivo central prover pontos de acesso a essas tendências monitoradas.

O trabalho tem caráter experimental por se tratar de um desenvolvimento de uma política de indexação que visa a melhorar um processo que já ocorre no OGI, principalmente reduzindo o seu nível de subjetividade. É importante salientar que essa característica exploratória objetiva entender se é possível aplicar uma nova forma de se pensar a indexação processada nesse ambiente, voltada para o ponto de vista do usuário das informações disponibilizadas pelo sistema.

As notícias são inseridas no OGI e categorizadas de acordo com parâmetros definidos por estudos anteriores realizados desde 2009. Essas categorias demonstram que as notícias que envolvem educação, eventos e mercado de trabalho, tanto em Gestão da Informação como em Gestão do Conhecimento.

Com a finalidade de aumentar a qualificação dessas notícias, demonstrando mais profundamente os temas e tendências que estão acontecendo nas duas áreas, foi decidido

aplicar as *tags*, seguindo a Política de Indexação definida para o OGI como foi definindo no item 4.4.

- a) Identificação da organização a qual está vinculado o sistema de indexação (contexto).

A política de indexação está sendo aplicada ao Observatório de Gestão da Informação, localizado na *web*, cuja autoria é de professores e alunos da Universidade Federal do Rio de Janeiro.

- b) Caracterização da clientela a que se destina o sistema (destinatário).

A clientela OGI é composta tanto de alunos, quanto de professores e público especializado em Organização da Informação e do Conhecimento, que podem acessar o sítio pela *internet*. Os objetivos do acesso são: ampliar o contato com o tema, como extensão dos aprendizados na disciplina Gestão da Informação e do Conhecimento, e obter novidades acerca dos temas.

- c) Cobertura de assuntos abordados pelo sistema.

Os assuntos abordados pelo OGI são educacionais, em nível de graduação e pós graduação, eventos, como congressos e encontros, e mercado de trabalho, analisando cargos e vagas. São incluídos assim textos sobre temas relacionados à Gestão da Informação e Gestão do Conhecimento, que serão monitorados e representados de acordo com a própria política de indexação.

- d) Determinação do(s) tipo(s) de documentos.

São selecionados gêneros jornalísticos, gêneros textuais emergentes com notícias e sites da *web*, selecionados para compor a base de dados do sistema.

- e) Determinação da(s) Língua(s).

Os documentos selecionados são escritos em Português do Brasil.

- f) Delimitação do período de tempo de publicação e cobertura dos assuntos

As notícias e documentos em análise foram publicadas nos seis primeiros meses do ano de 2011 no OGI.



g) Parâmetros para o Processo de indexação.

A indexação será seletiva obedecendo aos seguintes parâmetros.

Nível de exaustividade adotado pelo sistema:

Decidiu-se em atribuir de 3 a 5 termos, dependendo da riqueza dos assuntos apresentados nas notícias publicadas. Ou seja, como são textos dos mais variados tamanhos, a quantidade de conceitos relevantes selecionados pode variar. Portanto, indexação seletiva, na qual os assuntos apresentados no documento, potencialmente relevantes para o OGI, são identificados durante a análise conceitual e traduzidos em termos de indexação e podem ser recuperados pelo sistema. Como há notícias que não abordam esse número mínimo de assuntos, decidiu-se por não considerá-las e portanto não serão indexadas.

Nível de especificidade da linguagem de indexação:

O nível de especificidade da linguagem de indexação será o mais específico possível, pois a clientela é composta tanto de especialistas como de alunos de graduação e pós-graduação que já estão familiarizados com a terminologia das áreas em análise.

Capacidade de revocação e precisão do sistema:

A expectativa é de baixa revocação. O sistema, por conseguinte, tem por objetivo atingir alta precisão na recuperação de informações relevantes.

Linguagem de Indexação adotada pelo sistema:

Devido à diversidade de gêneros textuais e escritas adotadas, a indexação será preferencialmente derivativa, com o uso da linguagem natural das notícias. Porém, devido à necessidade de se monitorar tendências, alguns termos selecionados não estarão necessariamente no corpo textual da matéria. (Linguagem Híbrida)

Tipo de indexação segundo a coordenação de termos/conceitos:

Indexação pós-coordenada, segunda a qual a busca pode ser feita pela coordenação dos termos que são organizados na “nuvem de tags”. Portanto, os termos podem ser acessados, mas não podem ser combinados.

h) Avaliação do sistema e satisfação do usuário.

Os processos de indexação e recuperação da informação serão avaliados periodicamente a partir de consulta aos usuários, utilizando-se questionário *online*. Nestes questionários serão priorizados o nível de precisão do sistema e a relevância das informações recuperadas, do ponto de vista do usuário e domínio em análise.

## 6 RESULTADOS

Após a aplicação da Política de Indexação, nos seis primeiros meses de 2011 o resultado obtido foi considerável. Porém, faz-se necessário descrever todo o processo.

A política permitiu encontrar um mecanismo padronizado para a retirada dos termos, ao mesmo tempo em que possibilitou o aperfeiçoamento do processo, melhorando e ampliando a busca por novos temas relativos ao OGI. Esta política não visa a modificar a maneira como se indexa no OGI, mas sim apresentar melhoramentos ao processo.

Com a aplicação da política, nos seis primeiros meses de 2011, foi analisado um total de 142 notícias. Nessa análise, foram retiradas 33 etiquetas diferentes, como demonstrado na tabela 1. Quando a notícia apresentou pelo menos 3 assuntos relevantes, estes foram considerados e representados e assim constatou-se a frequência de 306 termos, sendo que alguns foram mais recorrentes e outros só apareceram uma única vez.

Já no ano anterior, quando não havia o estabelecimento de um padrão para o processo de indexação, foram analisadas 147 notícias, que geraram 19 etiquetas diferentes, com frequência 166, como apresentado na tabela 2. Como não havia um modelo estabelecido para a indexação, muitas notícias foram indexadas com apenas um termo, o que não representa especificamente o conteúdo temático das notícias.

É perceptível pela análise dos resultados que a qualidade do processo de indexação mudou de um período para o outro. Ao analisar apenas seis meses, podemos perceber uma considerável mudança na qualidade da indexação. Os resultados alcançados, tanto em 2010 como em 2011, apresentam quantidades equivalentes de notícias selecionadas; ou seja, 147 notícias em 2010 e 142 em 2011. Porém, ao analisar a quantidade de etiquetas na representação de

conceitos/assuntos abordados nessas notícias, pode-se notar a diferença: 19 etiquetas em 2010 e 33 etiquetas em 2011, apesar da restrição à quantidade de termos atribuídos por notícia (de 3 a cinco). Essa diferença se repete ao analisar a frequência de ocorrência dos termos: frequência de 166 termos em 2010 e de 306 em 2011.

Podemos perceber que, ao estabelecer um modelo de indexação, houve efetivamente um melhoramento no processo. A quantidade de assuntos monitorados aumentou no período de um ano em 73%, e a frequência também em 84%. A Política de Indexação permitiu que, em um mesmo período de tempo, seis meses, com uma quantidade de notícias bem próxima, (diferença de 5 notícias), fosse selecionado e indexado um maior número de assuntos, de forma mais representativa e enriquecedora, permitindo o aumento dos pontos de acesso na busca pelo sistema. Foram 14 novos assuntos monitorados em relação a 2010.

**Tabela 1 Contagem de etiquetas de 2011, primeiro semestre.**

<b>CONTAGEM DE ETIQUETAS 2011.1</b>			
<b>Administração</b>	25	<b>Finanças</b>	4
<b>Ambiental</b>	10	<b>Inovação</b>	15
<b>Aprendizagem</b>	5	<b>Inteligência Competitiva</b>	2
<b>Arquivologia</b>	5	<b>Marketing</b>	8
<b>Aspectos Sociais</b>	24	<b>Memória</b>	1
<b>Biblioteconomia</b>	5	<b>Mídia</b>	4
<b>Capital Humano</b>	1	<b>Museologia</b>	0
<b>Ciência da Informação</b>	6	<b>ONG</b>	4
<b>Comunicação</b>	9	<b>Psicologia</b>	3
<b>Criatividade</b>	2	<b>Recursos Humanos</b>	5
<b>Direito</b>	9	<b>Religião</b>	1
<b>Documentação</b>	7	<b>Saúde</b>	12
<b>Economia</b>	7	<b>Setor Privado</b>	42
<b>Empreendedorismo</b>	2	<b>Setor Público</b>	45
<b>Engenharia</b>	8	<b>Sustentabilidade</b>	8
<b>Esportes</b>	2	<b>Tecnologia da Informação</b>	24
<b>Ética</b>	1	<b>FREQUENCIA DOS TERMOS</b>	306

Os assuntos novos representados no ano de 2011.1 foram: Aprendizagem, Criatividade, Direito, Economia, Empreendedorismo, Esportes, Ética, Finanças, Marketing, Mídia, ONG, Psicologia, Recursos Humanos e Religião.

Tabela 2 Contagem de etiquetas de 2010, primeiro semestre.

<b>CONTAGEM DE ETIQUETAS 2010.1</b>			
Administração	18	Inovação	11
Ambiental	3	Inteligência Competitiva	6
Arquivologia	2	Memória	2
Aspectos Sociais	13	Museologia	2
Biblioteconomia	10	Saúde	4
Capital Humano	0	Setor Público	28
Ciência da Informação	9	Setor Privado	16
Comunicação	5	Sustentabilidade	4
Documentação	2	Tecnologia da Informação	23
Engenharia	8	<b>FREQUENCIA DOS TERMOS</b>	<b>166</b>

Ao se analisar os termos encontrados pode-se notar que são assuntos representados de forma não específica, como foi salientado ao apresentar o nível de especificidade, e com o intuito de apenas demonstrar atualidades e tendências nas áreas pesquisadas. Para os indexadores futuros dos repositórios disponibilizados no OGI, deve-se salientar que o processo é flexível, permitindo o surgimento de novos termos e assuntos que podem ser indexados demonstrando assim que as tendências sofrem transformações. Alguns assuntos estão mais presentes no decorrer de todos os meses monitorados, mas outros somente ocorrem algumas vezes, como demonstra a tabela 3.

Tabela 3 Contagem dos termos por meses em 2011.

<b>QUANTIDADE DE TERMOS INDEXADOS POR PERÍODO 2011.1</b>						
<b>Termos</b>	<b>Jan. 2011</b>	<b>Fev. 2011</b>	<b>Mar. 2011</b>	<b>Abril 2011</b>	<b>Mai. 2011</b>	<b>Jun. 2011</b>
Administração	3	8	6	6		2
Ambiental	4	2	1	0	1	2
Aprendizagem	2	1	2	0		0
Arquivologia	1	2	1	1		0
Aspectos Sociais	8	7	3	3	1	2
Biblioteconomia	2	1	1	1		0
Capital Humano	1	0	0	0		0
Ciência da Informação	2	1	1	1	1	0
Comunicação	3	3	1	2		0
Criatividade	2	0	0	0		0
Direito	6	0	1	1		1
Documentação	3	2	1	1		0

<b>Economia</b>	3	1	0	1		2
<b>Empreendedorismo</b>	0	0	1	1		0
<b>Engenharia</b>	3	0	4	1		0
<b>Esportes</b>	1	1	0	0		0
<b>Ética</b>	1	0	0	0		0
<b>Finanças</b>	0	1	2	0		1
<b>Inovação</b>	6	2	5	1		1
<b>Inteligência Competitiva</b>	1	0	0	1		0
<b>Marketing</b>	5	2	1	0		0
<b>Memória</b>	0	0	0	0	1	0
<b>Mídia</b>	2	1	0	1		0
<b>Museologia</b>	0	0	0	0		0
<b>ONG</b>	1	0	0	0	1	2
<b>Psicologia</b>	1	0	0	1		1
<b>Recursos Humanos</b>	2	2	0	0		1
<b>Religião</b>	1	0	0	0		0
<b>Saúde</b>	5	5	1	1		0
<b>Setor Privado</b>	19	6	9	5		3
<b>Setor Público</b>	11	13	6	8	1	6
<b>Sustentabilidade</b>	4	1	3	0		0
<b>Tecnologia da Informação</b>	6	5	5	4	1	3
<b>TOTAL</b>	<b>109</b>	<b>67</b>	<b>55</b>	<b>41</b>	<b>7</b>	<b>27</b>

Os termos monitorados ou etiquetas não necessariamente se encontram no corpo textual das notícias, e também não são baseados em um vocabulário controlado. Mas faz-se necessário utilizar de uma mecânica que permita descobrir qual termo será adotado na indexação a partir de certos assuntos presentes. Como o OGI visa a monitorar somente as tendências, decidiu-se aplicar um modelo de extração das etiquetas através das seguintes características:

- 1) Termos como Administração, Arquivologia, Biblioteconomia, Ciência da Informação, Comunicação, Direito, Documentação, Economia, Engenharia, Museologia, Psicologia e Tecnologia da Informação, são termos encontrados no corpo textual das notícias, portanto seguem o critério da indexação por linguagem natural dos textos. E indicam as abordagens disciplinares apresentadas nas notícias, como notícias de universidades, faculdades, cursos, especialistas etc.
- 2) Os termos Setor Público, Setor Privado e ONG, não necessariamente são encontrados na parte textual das notícias, mas são extraídos de acordo com as

necessidade de se monitorar onde estão surgindo as tendências. Para aplicar os termos seguiu-se o seguinte modelo.

- a) Setor Público está relacionado à notícias que falam sobre: Políticas públicas, programas governamentais, secretárias municipais, agências do governo, departamentos, ministérios etc.
- b) Setor Privado, a notícia está apresentando os assuntos: Comércio, competição empresarial, corporativismo, negócios entre empresas etc.
- c) ONG, indica as notícias que abordam sobre iniciativas que não englobam os dois setores anteriormente monitorados: Organizações não governamentais.

3) Para o último grupo de termos, sendo eles: Ambiental, Aprendizagem, Aspectos Sociais, Capital Humano, Criatividade, Empreendedorismo, Esportes, Finanças, Inovação, Inteligência Competitiva, Marketing, Memória, Mídia, Recursos Humanos, Religião, Saúde e Sustentabilidade, foram considerados os assuntos que demonstram aspectos ligados às finalidades que a notícia apresenta. Estes termos podem ser encontrados nas notícias ou foram interpretados de acordo com o objetivo de se monitorar alguma atividade, sendo assim não foram encontrados no corpo textual da matéria. Os termos interpretados são:

- a) Ambiental, relacionado a notícias que falam sobre: meio ambiente, mata atlântica, natureza, energias renováveis, poluição, resíduos, reciclagem etc.
- b) Aspectos Sociais, originado em notícias que apresentam os assuntos: Desenvolvimento social, responsabilidade social, segurança familiar, comunidades, igualdade social, combate à fome, qualidade de vida etc.
- c) Mídia, a partir de assuntos relacionados aos meios de comunicação, como: Rádio, TV, jornal, revistas, *internet* etc.
- d) Saúde, para os assuntos: Alimentação, nutrição, hospitais, etc.

É importante entender que essas 3 características não necessariamente devem estar presentes nas matérias, algumas notícias irão apresentar mais características disciplinares do que elementos relacionados aos setores ou finalidade. Com o intuito de entender melhor como essas notícias estão surgindo, foi desenvolvida uma tabela que demonstra as hierarquias já apresentadas nestas 3 características.

**Tabela 4 Hierarquia dos termos.**

<b>HIERARQUIA DOS TIPOS DE TERMOS</b>		
<b>ABORDAGEM</b>	<b>ONDE?</b>	<b>PARA QUÊ?</b>
Administração	Setor Privado	Ambiental
Arquivologia	Setor Público	Aprendizagem
Biblioteconomia	ONG	Aspectos Sociais
Ciência da Informação		Capital Humano
Comunicação		Criatividade
Direito		Empreendedorismo
Documentação		Esportes
Economia		Ética
Engenharia		Finanças
Museologia		Inovação
Psicologia		Inteligência Competitiva
Tecnologia da Informação		Marketing
		Memória
		Mídia
		Recursos Humanos
		Religião
		Saúde
		Sustentabilidade

Através desta classificação pode-se entender melhor a finalidade de cada termo na indexação no OGI. Para Abordagem, são colocados os termos que apresentam os aspectos disciplinares das notícias. Em Onde?, ficam os termos que têm origem nos setores mencionados e Para Quê? demonstra a finalidade do uso de cada um dos respectivos termos na notícia.

## 7 CONSIDERAÇÕES

Após a aplicação da política de indexação no OGI e análise dos resultados, conclui-se que os objetivos foram alcançados. No entanto, alguns pontos devem ser salientados.

A política de indexação aplicada ao OGI teve por objetivo melhorar o processo de representação da informação que já existia. Portanto não faz parte do interesse do trabalho modificar a forma como se indexava, mas sim aperfeiçoar o processo com o estabelecimento de parâmetros e elementos definidos previamente na política de indexação, com o propósito de tentar reduzir as parcialidades e subjetividades inerentes ao indexador ao analisar as notícias.

Este modelo de indexação tem também por objetivo servir como padrão na indexação das informações capturadas por futuros indexadores, que poderão fazer uso da política aqui delineada, visando a dar continuidade à tarefa, de forma padronizada e de acordo com os objetivos do OGI.

É importante considerar que o modelo de indexação apresentado no trabalho visa ao estabelecimento de parâmetros para a atividade em si. A aplicação deste modelo deixou algumas lacunas para o melhoramento da indexação em desdobramentos futuros, que visem também a ampliar a especificidades dos assuntos monitorados, a fim de se obter uma análise cada vez mais aprofundada do tema pesquisado. O nível de especificidade pode ser intensificado em trabalhos futuros justamente pelo fato de ter sido constatado que alguns termos, principalmente os atribuídos, originam-se em assuntos mais específicos nas notícias.

Por fim, o projeto OGI tem por característica apresentar um modelo de monitoramento de assunto. Esse modelo pode ser aplicado a qualquer área de interesse, seguindo a sua metodologia.



## REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: informação e documentação: referências: elaboração. Rio de Janeiro, 2002.

\_\_\_\_\_. **NBR 10520**: informação e documentação: citações em documentos: apresentação. Rio de Janeiro, 2002.

\_\_\_\_\_. **NBR 14724**: informação e documentação: trabalhos acadêmicos: apresentação. Rio de Janeiro, 2005.

FEITOSA, A. **Organização da informação na web**: das tags à web semântica. Brasília: Thesaurus, 2006.

FUJITA, M. S. L. A identificação de conceitos no processo de análise de assunto para indexação. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 1, n. 1, p. 60-90, jul./dez. 2003.

\_\_\_\_\_, M. S. L. A Leitura Documentária na perspectiva de suas variáveis: leitor, texto, contexto. **Revista de Ciência da Informação**, Brasília, DF, v. 5, n. 4, ago. 2004.

\_\_\_\_\_, M. S. L.; RUBI, M. P. O ensino de procedimentos de política de indexação na perspectiva do conhecimento organizacional: uma proposta de programa para a educação à distância do bibliotecário. **Perspectivas em ciência da informação**, Belo Horizonte, v.11 n.1, p. 48-66, jan./abr. 2006.

\_\_\_\_\_, M. S. L.; SILVA, M. R. A prática de indexação: análise da evolução de tendências teóricas e metodológicas. **Transinformação**, Campinas, v. 16, n. 2, p. 133-161, maio/ago. 2004.

GUEDES, V. L. S. Indexação e recuperação da Informação: princípios, conceitos e considerações. In: **IX ENCONTRO NACIONAL DE ENSINO E PESQUISA EM INFORMAÇÃO**, 9, 2009, Salvador. Anais... Salvador: UFBA, 2009.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. Tradução: Antônio Agenor Briquet de Lemos. 2. ed. ver. atual. Brasília, DF: Briquet de Lemos, 2004.

LE COADIC, Y. **A ciência da informação**. Tradução: Maria Yêda F. S. de Filgueiras Gomes. 2. Ed. Brasília, DF: Briquet de Lemos, 2004.

MALIN, A. M. B. et al. **Observatório de Gestão da Informação**. Rio de Janeiro, 2010. Disponível em: <<http://ogimonitorandoagora.wordpress.com/>>. Acesso em: 5 jul. 2011

MATIAS-PEREIRA, J. **Manual de metodologia da pesquisa científica**. São Paulo: Atlas, 2007.

PAZ, D. M. S. A teoria de Gough e o modelo ascendente de leitura. **Linguagens & Cidadania**, Santa Maria, v. 11, p. 12-22, 2006.

ROCHA, R. P. Metadados, web semântica, categorização automática: combinando esforços humanos e computacionais para a descoberta e uso dos recursos da web. **Questão**, Porto Alegre, v. 10, n. 1, p. 109-121, jan./jun. 2004.

UNITED NATIONS INTERNATIONAL SCIENTIFIC INFORMATION SYSTEM.  
**Indexing principies**. Tradução Maria Cristina Mello Ferreira Pinto. Paris, UNESCO, 1976, 11 p.